**Remarks**

The specification and claims have been amended to correctly recite "SEQ ID NO:" rather than "SEQ ID" as required by 37 CFR 1.821(d).

Claims 16 and 20, which are directed to polynucleotides, are amended to delete SEQ ID NOS:103-104 which are amino acid sequences. Claim 26, which is directed to a polypeptide, is amended to delete SEQ ID NOS:115, 110, and 111, which are nucleotide sequences. The amendments add no new matter.

Respectfully submitted,
BANNER & WITCOFF, LTD.

Date:

Dale H. Hoscheit
Registration No. 19,090

1001 G Street, N.W., Eleventh Floor
Washington, D.C. 20001-4597
(202) 508-9100

Appendix 1.Version of Amended Claims and Paragraphs with Markings to Show Changes Made

Bracketed language is present in the original. Deletions are shown in Strikethroughs.

4. (Amended) The method of any preceding claim, wherein the expression product is a RNA having the following formula: $N_1$-$N_2$-$N_3$-$N_4$-$N_5$-polyA, wherein: $N_1$ has at least x% sequence identity to SEQ ID NO:155; $N_2$ has at least x% sequence identity to SEQ ID NO:156; $N_3$ has at least x% sequence identity to SEQ ID NO:6; $N_4$ comprises any RNA sequence; $N_5$ has at least x% sequence identity to SEQ ID NO:5; and at least one of $N_1$ or $N_5$ is present, but $N_2$, $N_3$, $N_4$ and polyA are optional.

16. (Amended) Isolated polynucleotide of claim 16, comprising one of SEQ ID NOS:7-39, SEQ ID NOS:44-45, SEQ ID NOS:59-91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NOS: ~~99-105~~ 99-102, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NOS:110-145, SEQ ID NOS:150-157, or SEQ ID NOS:161-225.

20. (Amended) Isolated polynucleotide of claim 21, wherein B comprises a fragment of 7 or more nucleotides from one or more of SEQ ID NOS:7-39, 44-45, 59-91, 93, 95, 97, ~~99-105,~~ 99-102, 105, 107, 110-145, 150-157, and 161-225.

25. (Amended) An isolated polypeptide for use in the diagnosis of prostate cancer, the polypeptide comprising: (a) an amino acid sequence selected from the group consisting of SEQ ID NOS:109, 146, 147, 148 and 149; (b) a fragment of at least 7 amino acids of (a); or (c) a polypeptide sequence having at least 50% identity to (a).

26. (Amended) An isolated polypeptide having formula NH2-A-B-C-COOH, wherein: A is a polypeptide sequence consisting of *a* amino acids; C is a polypeptide sequence consisting of *c* amino acids; B is a polypeptide sequence consisting of a fragment of at least 5 amino acids of an

amino acid sequence selected from the group consisting of SEQ ID NOS:146, 147, 148, 149, and ~~115,~~ 109 ~~, 110 and 111~~; and said polypeptide is not a fragment of polypeptide sequence SEQ ID NO:146, 147, 148, 149, or ~~115,~~ 109 ~~, 110 or 111~~; and wherein $a+c \geq 1$.

(1) Page 3, line 22, to page 4, line 15:

A sequence which has at least 75% identity to SEQ ID NO:155 (e.g. 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID NO:155 (e.g. 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, etc., contiguous nucleotides) of SEQ ID NO:155; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, etc., contiguous nucleotides) of SEQ ID NO:155 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. This sequence will typically be at the 5' end of the RNA. SEQ ID NO:155 is the nucleotide sequence of the start of R region in the LTR of the 'ERVK6' HML-2 virus [ref. 1]. This portion of the R region is found in all full-length HML-2 transcripts.

(2) Page 4, line16, to page 5, line 9:

A downstream region comprising a sequence which has at least 75% sequence identity to SEQ ID NO:156 (e.g. 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%,

89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID NO:156 (e.g. 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, etc., contiguous nucleotides) of SEQ ID NO:156; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, etc., contiguous nucleotides) of SEQ ID NO:156 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. SEQ ID NO:156 is the nucleotide sequence of the RU5 region downstream of SEQ ID NO:155 in the ERVK6 LTR. This region is found in full-length HML-2 transcripts, but may not be present in all mRNAs transcribed from a HML-2 LTR promoter.

(3) Page 5, line 10 to line 31:

A downstream region comprising a sequence which has at least 75% sequence identity to SEQ ID NO:6 (e.g. 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID NO:6 (e.g. 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,

88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, etc., contiguous nucleotides) of SEQ ID NO:6; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, etc., contiguous nucleotides) of SEQ ID NO:6 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. SEQ ID NO:6 is the nucleotide sequence of the region of the ERVK6 virus between the U5 region and the first 5' splice site. This region is found in full-length HML-2 transcripts, but has been lost by some variants and, like region 2 above, may not be present in all mRNAs transcribed from a HML-2 LTR promoter.

(4) Page 6, lines 6-31:

A downstream region comprising a sequence which has at least 75% sequence identity to SEQ ID NO:5 (e.g. 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID NO:5 (e.g. 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185,

190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc., contiguous nucleotides) of SEQ ID NO:5; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. . 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc., contiguous nucleotides) of SEQ ID NO:5 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. SEQ ID NO:5 is the nucleotide sequence of the U3R region in the 3' end of ERVK6. This sequence will typically be near the 3' end of the RNA, immediately preceding any polyA tail.

(5) Page 7, line18, to page 8, line 20:

— N1 has at least 75% sequence identity to SEQ ID NO:155; or has at least 50% identity to SEQ ID NO:155 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:155; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:155 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level;

— N2 has at least 75% sequence identity to SEQ ID NO:156; or has at least 50% identity to SEQ ID NO:156 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:156; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:156 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level;

— N3 has at least 75% sequence identity to SEQ ID NO:6; or has at least 50% identity to SEQ ID NO:6 and is expressed at least 1.5 fold higher relative to expression in a

normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:6; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:6 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level;

— N4 comprises any RNA sequence;

— N5 has at least 75% sequence identity to SEQ ID NO:5; or has at least 50% identity to SEQ ID NO:5 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:5; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID NO:5 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; and

— at least one of $N_1$, $N_2$, $N_3$, $N_4$ or $N_5$ is present, but polyA is optional.

(6) Page 9, lines 5-12:

Where diagnosis is based on mRNA detection, the method of the invention preferably comprises an initial step of: (a) extracting RNA (*e.g.* mRNA) from a patient sample; (b) removing DNA from a patient sample without removing mRNA; and/or (c) removing or disrupting DNA which comprises SEQ ID NO:4, but not RNA which comprises SEQ ID NO:4, from a patient sample. This is necessary because the genomes of both normal and cancerous prostate cells contain multiple PCAV DNA templates, whereas increased PCA-mRNA levels are only found in cancerous cells. As an alternative, a RNA-specific assay can be used which is not affected by the presence of homologous DNA.

(7) Page 9, lines 18-21:

Methods for removing DNA, but not RNA, comprising PCA-mRNA sequences will use a reagent which is specific to a sequence within a PCA-mRNA *e.g.* a restriction enzyme which recognizes a DNA sequence within SEQ ID NO:4, but which does not cleave the corresponding RNA sequence.

(8) Page 16, lines 7-9:

Examples of gag nucleotide sequences are: SEQ ~~IDs~~ ID NOS:7, 8, 9 & 11 [HERV-

K(CH)]; SEQ ID NO:85 [HERV-K108]; SEQ ID NO:91 [HERV-K(C7)]; SEQ ID NO:97 [HERV-K(II)]; SEQ ID NO:102 [HERV-K10].

(9) Page 16, lines 10-12:

Examples of gag polypeptide sequences are: SEQ ~~IDs~~ ID NOS:46, 47, 48, 49, 56 & 57 [HERV-K(CH)]; SEQ ID NO:92 [HERV-K(C7)]; SEQ ID NO:98 [HERV-K(II)]; SEQ ~~IDs~~ ID NOS:103 & 104 [HERV-K10] ; SEQ ID NO:146 ['ERVK6'].

(10) Page 16, lines 17-18:

Examples of prt nucleotide sequences are: SEQ ID NO:86 [HERV-K(108)]; SEQ ID NO:99 [HERV-K(II)]; SEQ ID NO:105 [HERV-K10].

(11) Page 16, lines 19-20:

Examples of prt polypeptide sequences are: SEQ ID NO:106 [HERV-K10]; SEQ ID NO:147 ['ERVK6'].

(12) Page 16, lines 24-25:

Examples of pol nucleotide sequences are: SEQ ID NO:87 [HERV-K(108)]; SEQ ID NO:93 [HERV-K(C7)]; SEQ ID NO:100 [HERV-K(II)]; SEQ ID NO:107 [HERV-K10].

(13) Page 16, lines 26-27:

Examples of pol polypeptide sequences are: SEQ ID NO:94 [HERV-K(C7)]; SEQ ID NO:108 [HERV-K10]; SEQ ID NO:148 ['ERVK6'].

(14) Page 17, lines 4-5:

Examples of env nucleotide sequences are: SEQ ID NO:88 [HERV-K(108)]; SEQ ID NO:95 [HERV-K(C7)]; SEQ ID NO:101 [HERV-K(II)]; SEQ ID NO:107 [HERV-K10].

(15) Page 17, lines 6-7:

Examples of env polypeptide sequences are: SEQ ID NO:96 [HERV-K(C7)]; SEQ ID NO:108 [HERV-K10] ; SEQ ID NO:149 ['ERVK6'].

(16) Page 17, line 13:

Examples of cORF nucleotide sequences are: SEQ ID NO:89 and SEQ ID NO:90 [HERV-K(108)]

(17) Page 17, line 14:

Examples of cORF polypeptide sequences are SEQ ID NO:109.

(18) Page 19, lines 4-8:

The invention provides an isolated polypeptide comprising: (a) an amino acid sequence selected from the group consisting of SEQ ~~IDs~~ ID NOS:109 (cORF), 146 (gag), 147 (prt), 148 (pol), 149 (env); (b) a fragment of at least *x* amino acids of (a); or (c) a polypeptide sequence having at least *s*% identity to (a). These polypeptides include variants (*e.g.* allelic variants, homologs, orthologs, mutants *etc.*).

(19) Page 19, lines 14-18:

The invention also provides an isolated polypeptide having formula NH2-A-B-C-COOH, wherein: A is a polypeptide sequence consisting of *a* amino acids; C is a polypeptide sequence consisting of *c* amino acids; B is a polypeptide sequence consisting of a fragment of *b* amino acids of an amino acid sequence selected from the group consisting of SEQ ~~IDs~~ ID NOS:109, 146, 147, 148, 149; and said polypeptide is not a fragment of polypeptide sequence SEQ ID NO:109, 146, 147, 148 or 149.

(20) Page 19, line 27, to page 20, line 2:

The amino acid sequence of -A- typically shares less than *n*% sequence identity to the *a* amino acids which are N-terminal of sequence -B- in SEQ ID NO:109, 146, 147, 148 or 149 and the amino acid sequence of -C- typically shares less than *n*% sequence identity to the *c* amino acids which are C-terminal of sequence -B- in SEQ ID NO:109, 146, 147, 148 or 149. The value of *n* is generally 60 or less (*e.g.* 50, 40, 30, 20, 10 or less).

(21) Page 20, lines 3-8:

The fragment of (b) may comprise a T-cell or, preferably, a B-cell epitope of SEQ ID NO:109, 146, 147, 148 or 149. T- and B-cell epitopes can be identified empirically (*e.g.* using the PEPSCAN method [19, 20] or similar methods), or they can be predicted (*e.g.* using the Jameson-Wolf antigenic index [21], matrix-based approaches [22], TEPITOPE [23], neural networks [24], OptiMer & EpiMer [25, 26], ADEPT [27], Tsites [28], hydrophilicity [29], antigenic index [30] or the methods disclosed in reference 31 *etc.*).

(22) Page 37, lines 10-15:

The invention is based on the finding that HML-2 mRNA expression is up-regulated in prostate tumors. Because HML-2 is a well-recognized family, the skilled person will be able to determine without difficulty whether any particular endogenous retroviruses is or is not a HML-

2. Preferred members of the HML-2 family for use in accordance with the present invention are those whose proviral genome has an LTR which has at least 75% sequence identity to SEQ ID NO:150 (the LTR sequence from HML-2.HOM [1]). Example LTRs include SEQ ~~IDs~~ ID NOS:151-154.

(23) Page 37, lines 20-27:

Sequences from HERV-K(CH) are shown in SEQ ~~IDs~~ ID NOS:14-39 and have been deposited with the ATCC (see Table 7). The skilled person will be able to classify any further HERV as HERV-K(CH) or not based on sequence identity to these HERV-K(CH) polynucleotides. Preferably such a comparison is to one or more, or all, of the polynucleotide sequences disclosed herein or of the polynucleotide inserts in the ATCC-deposited isolates. Alternatively, the skilled artisan can determine the sequence identity based on a comparison to any one or more, or all, of the sequences in SEQ ~~IDs~~ ID NOS:7-10 and SEQ ~~IDs~~ ID NOS:14-39 taking into consideration the spontaneous mutation rate associated with retroviral replication. Thus, it will be apparent when the differences in the sequences are consistent with a HERV-K(CH) isolate or consistent with another HERV.

(24) Page 38, lines 7-10:

The invention provides an isolated polynucleotide comprising: (a) the nucleotide sequence of any of SEQ ~~IDs~~ ID NOS:7-10; (b) the nucleotide sequence of any of SEQ ~~IDs~~ ID NOS:27-39; (c) the complement of a nucleotide sequence of any of SEQ ~~IDs~~ ID NOS:7-10; or (d) the complement of the nucleotide sequence of any of SEQ ~~IDs~~ ID NOS:27-39.

(25) Page 38, lines 12-15:

The invention also provides an isolated polynucleotide comprising a fragment of: (a) a nucleotide sequence shown in SEQ ~~IDs~~ ID NOS:7-10; (b) the nucleotide sequence shown in any of SEQ ~~IDs~~ ID NOS:27-39; (c) the complement of a nucleotide sequence shown in SEQ ~~IDs~~ ID NOS:7-10; or (d) the complement of the nucleotide sequence shown in any of SEQ ~~IDs~~ ID NOS:27-39.

(26) Page 38, lines 22-27:

The fragment is preferably neither one of the following sequences nor a fragment of one of the following sequences: (i) the nucleotide sequence shown in SEQ ID NO:42; (ii) the nucleotide sequence shown in SEQ ID NO:43; (iii) the nucleotide sequence shown in SEQ ID

NO:44; (iv) the nucleotide sequence shown in SEQ ID NO:45; (v) a known polynucleotide; or (vi) a polynucleotide known as of 7th December 2000 (*e.g.* a polynucleotide available in a public database such as GenBank of GeneSeq before 7th December 2000).

(27) Page 39, lines 14-17:

Preferred fragments (*e.g.* for the identification of HERV-K(CH) polynucleotides associated with cancer) which do not correspond identically in their entirety to any portion of the sequence(s) shown in SEQ ~~IDs~~ ID NOS:42-45 are: SEQ ID NO:59 (from gag region), SEQ ~~IDs~~ ID NOS:60-70 (from pol region) and SEQ ~~IDs~~ ID NOS:71-82 (from 3' pol region).

(28) Page 39, lines 17-21:

Preferred fragments (*e.g.* for the simultaneous identification of HERV-K(CH) polynucleotides, HERV-KII polynucleotides and/or HERV-K10 polynucleotides) which do correspond identically in their entirety to any portion of the sequence(s) shown in SEQ ~~IDs~~ ID NOS:44 & 45 are SEQ ~~IDs~~ ID NOS:83 & 84 (from gag region).

(29) Page 39, line 27, to page 40, line 9:

The invention also provides an isolated polynucleotide comprising (a) a segment that is a fragment of the sequence shown in SEQ ~~IDs~~ ID NOS:7-10 or SEQ ~~IDs~~ ID NOS:27-39, wherein (i) said fragment is at least 10 nucleotides in length and (ii) corresponds identically in its entirety to a portion of SEQ ID NO:44 and/or 45; and, optionally, (b) one or more segments flanking the segment defined in (a), wherein the presence of said optional segment(s) causes said polynucleotide to not correspond identically to any portion of a sequence shown in SEQ ~~IDs~~ ID NOS:7- 10 or SEQ ~~IDs~~ ID NOS:27-39. In some embodiments, the optional flanking segments share less than 40% sequence identity to the nucleic acid sequences shown in SEQ ~~IDs~~ ID NOS:7-10, SEQ ID NO:44 and/or SEQ ID NO:45. In other embodiments, the optional flanking segments have no contiguous sequence of 10, 12, 15 or 20 nucleotides in common with SEQ ~~IDs~~ ID NOS:7-10, SEQ ID NO:44 and/or SEQ ID NO:45. In yet other embodiments, the optional flanking segment is not present. In further embodiments, a fragment of the polynucleotide sequence is up to at least 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000, or 1500 nucleotides in length.

(30) Page 40, lines 10-19:

The invention also provides an isolated polynucleotide having formula 5'-A-B-C-3',

wherein: A is a nucleotide sequence consisting of *a* nucleotides; B is a nucleotide sequence consisting of a fragment of *b* nucleotides from (i) the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10, (ii) the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39, (iii) the complement of the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10, or (iv) the complement of the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39; C is a nucleotide sequence consisting of *c* nucleotides; and wherein said polynucleotide is not a fragment of (i) the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10, (ii) the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39, (iii) the complement of the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10, or (iv) the complement of the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39.

(31) Page 41, lines 2-6:

The invention provides a polynucleotide having at least *s*% identity to: (a) SEQ ~~ID~~ ID NOS:7-10; (b) a fragment of *x* nucleotides of SEQ ~~ID~~ ID NOS:7-10; (c) SEQ ~~ID~~ ID NOS:11-13; (b) a fragment of *x* nucleotides of SEQ ~~ID~~ ID NOS:11-13. The value of *s* is at least 50 (*e.g.* at least 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.5, 99.9 *etc.*). The value of *x* is at least 7 (*e.g.* 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 *etc.*).

(32) Page 41, lines 9-14:

Variants can be identified by hybridization of putative variants with the polynucleotide sequences disclosed in SEQ ~~ID~~ ID NOS:14-39 herein, preferably by hybridization under stringent conditions. For example, by using appropriate wash conditions, variants can be identified where the allelic variant exhibits at most about 25-30% base pair (bp) mismatches relative to the selected polynucleotide probe. In general, allelic variants contain 15-25% bp mismatches, and can contain as little as even 5-15%, or 2-5%, or 1-2% bp mismatches, as well as a single bp mismatch.

(33) Page 41, line 25, to page 42, line 12:

A preferred HERV-K(CH) isolate is an isolate sequence which is shown in SEQ ~~ID~~ ID NOS:7-10. Another preferred class of HERV-K(CH) isolates are those having a nucleotide sequence identity of at least 90%, preferably at least 95% to the 3' polymerase region shown in SEQ ID NO:13 which relates to integrase, as measured by the alignment program GCG Gap

(Suite Version 10.1) using the default parameters: open gap = 3 and extend gap = 1. Another preferred class of HERV-K(CH) isolates are those having a nucleotide sequence identity of at least 98%, more preferably at least 99% to the 5' polymerase region shown in SEQ ID NO:12 which relates to reverse transcriptase, as measured by the alignment program GCG Gap (Suite Version 10.1) using the default parameters: open gap = 3 and extend gap = 1. Another typical classification of the relationship of retroviruses is based on the amino acid sequence similarities in the reverse transcriptase protein. Thus, an even more preferred class of HERV-K(CH) isolates are those having an amino acid sequence identity of at least 90%, more preferably 95% to the 5' polymerase region encoded by the nucleotide sequence shown in SEQ ID NO:12, as determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. Thus, these prostate cancer-associated polynucleotide sequences define a class of human endogenous retroviruses, designated herein as HERV-K(CH), whose members comprise variations which, without wanted to be bound by theory, may be due to the presence of polymorphisms or allelic variations.

(34) Page 42, lines 14-25:

The invention provides an isolated polynucleotide comprising a polynucleotide that selectively hybridizes, relative to a known polynucleotide, to: (a) the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10; (b) the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39; (c) the complement of the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10; (d) the complement of the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39; (e) a fragment of the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10; (f) a fragment of the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39; (g) the complement of a fragment of the nucleotide sequence shown in SEQ ~~ID~~ ID NOS:7-10; (h) the complement of a fragment of the nucleotide sequence shown in any of SEQ ~~ID~~ ID NOS:27-39; (j) a nucleotide sequence shown in SEQ ~~ID~~ ID NOS:14-39; or (k) polynucleotides found in ATCC deposits having ATCC accession numbers given in Table 7. The fragment of (e), (f), (g) or (h) is preferably at least *x* nucleotides in length, wherein *x* is as defined in H.1.2 above, and is preferably not one of the sequences (i), (ii), (iii), (iv), (v) or (vi) as defined H.1.2 above.

(35) Page 43, lines 2-7:

The invention also provides an isolated polynucleotide comprising: (a) a HERV-K(CH)

cDNA insert as deposited at the ATCC and having an ATCC accession number given in Table 7; (b) a HERV-K(CH) sequence as shown in any one of SEQ ~~IDs~~ ID NOS:14-26; (c) a HERV-K(CH) sequence as shown in any one of SEQ ~~IDs~~ ID NOS:27-39; or (d) a fragment of (a), (b) or (c). The fragment of (d) is preferably at least *x* nucleotides in length, wherein *x* is at least 7 (*e.g.* at least 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 *etc.*).

(36) Page 41, lines 9-22:

Preferred polynucleotides of the invention are those having a sequence set forth in any one of the polynucleotide sequences SEQ ~~IDs~~ ID NOS:7-10 and SEQ ~~IDs~~ ID NOS:14-39 provided herein; polynucleotides obtained from the biological materials described herein, in particular, polynucleotide sequences present in the isolates deposited with the ATCC and having ATCC accession numbers given in Table 7 or other biological sources (particularly human sources) or by hybridization to the above mentioned sequences under stringent conditions (particularly conditions of high stringency); genes corresponding to the provided polynucleotides; variants of the provided polynucleotides and their corresponding genes particularly those variants that retain a biological activity of the encoded gene product (*e.g.* a biological activity ascribed to a gene product corresponding to the provided polynucleotides as a result of the assignment of the gene product to a protein family(ies) and/or identification of a functional domain present in the gene product). Other polynucleotides and polynucleotide compositions contemplated by and within the scope of the present invention will be readily apparent to one of ordinary skill in the art when provided with the disclosure here.

(37) Page 45, lines 9-11:

The invention provides an isolated polypeptide: (a) encoded within a HERV-K(CH) open reading frame; (b) encoded by a polynucleotide shown in SEQ ID NO:11, 12 or 13; or (c) comprising an amino acid sequence as shown in any one of SEQ ~~IDs~~ ID NOS:46-49, 50-55, 56-57 or 58.

(38) Page 45, lines 12-18:

Deduced polypeptides encoded by the HERV-K(CH) polynucleotides of the invention include the gag translations shown in SEQ IDS 46-49 and the 3' pol translations shown in SEQ ~~IDs~~ ID NOS:50-55. A polypeptide sequence encoded by the polynucleotide having the sequence

shown in SEQ ID NO:15 is provided in SEQ ID NO:56; a polypeptide sequence encoded by the polynucleotide having the sequence shown in SEQ ID NO:14, is shown in SEQ ID NO:57. A consensus 3' pol polypeptide sequence encoded by the polynucleotides having the sequence shown in SEQ ~~ID~~s ID NOS:21-27, inclusive, is provided in SEQ ID NO:58.

(39) Page 45, lines 19-25:

The polypeptides encompassed by the present invention include those encoded by polynucleotides of the invention, *e.g.* SEQ ~~ID~~s ID NOS:7-10 and SEQ ~~ID~~s ID NOS:14-39, as well as polynucleotides deposited with the ATCC as disclosed herein, as well as nucleic acids that, by virtue of the degeneracy of the genetic code, are not identical in sequence to the disclosed polynucleotides and encode the polypeptides. Thus, the invention includes within its scope a polypeptide encoded by a polynucleotide having the sequence of any one of the polynucleotide sequences provided herein, or a variant thereof.

(40) Page 46, lines 17-20:

The invention provides an isolated polypeptide comprising a fragment of: (a) a polypeptide sequence encoded within a HERV-K(CH) open reading frame; (b) a polypeptide sequence encoded by a polynucleotide shown in SEQ ID NO:11, 12 or 13; or (c) an amino acid sequence as shown in any one of SEQ ~~ID~~s ID NOS:46-49, 50-55, 56-57 or 58.

(41) Page 46, lines 25-26:

The fragment may include an epitope *e.g.* an epitope of the amino acid sequence shown in SEQ ~~ID~~s ID NOS:56, 57 or 58.

(42) Page 46, line 27, to page 47, line 10:

SEQ ~~ID~~s ID NOS:46-49 provide a translation of the HERV-K(CH) polynucleotides having a sequence shown in SEQ ~~ID~~s ID NOS:14, 15, 16 and 40 (the sequence of SEQ ID NO:40 is from a polynucleotide found in a normal prostate library) corresponding to polynucleotides encoding the gag region. SEQ ~~ID~~s ID NOS:50-55 provide a translation of the HERV-K(CH) polynucleotides having a sequence shown in SEQ ~~ID~~s ID NOS:21-26, inclusive, corresponding to the 3' region of pol. SEQ ~~ID~~s ID NOS:56 & 57 provide translations of the HERV-K(CH) polynucleotide of SEQ ID NO:15 and SEQ ID NO:14, respectively. SEQ ID NO:58 provides a consensus translation of the polynucleotide from the 3' pol region (SEQ ~~ID~~s ID NOS:21-26, inclusive). Encompassed with the present invention are polypeptide fragments,

such as, epitopes, of at least 5 amino acids, at least 6 amino acids, at least 8 amino acids, at least 10 amino acids, at least 11 amino acids, at least 12 amino acids, at least 13 amino acids, at least 14 amino acids and at least 15 amino acids of the translations shown in SEQ ~~IDs~~ ID NOS:46-49 and 50-55. In a preferred embodiment, the HERV-K(CH) epitopes of the amino acid sequence as shown in SEQ ~~IDs~~ ID NOS:56-58 were determined by the Jameson-Wolf antigenic index [21].

(43) Page 47, lines 11-18:

The following regions in 3' pol (SEQ ID NO:58) were determined to be antigenic by Jameson-Wolf algorithm: amino acids: 1-10; 15-35; 45-55; 60-85; 100-115; 125-140; 170-190; 195-215; 230-268. Additional epitope-containing fragments include amino acids 1-8; 2-10; 1-15; 5-15; 7-15; 10-20; 12-20; 15-23; 20-28; 28-35; 15-30; 15-40; 20-30; 45-52; 48-55; 60-68; 60-70; 65-73; 70-78; 75-83; 70-80; 65-75; 68-75; 75-85; 78-85; 65-85; 60-75; 100-108; 103-110; 105-113; 108-115; 125-133; 128-135; 132-140; 170-178; 175-182; 180-187; 182-190; 195-202; 200-208; 205-212; 208-215; 230-237; 235-242; 240-247; 245-252; 250-257; 255-262; 260-268; 230-250; 235-255; 240-260; 245-268; 230-245; 235-245; 235-250; 240-255; 245-260; 250-268; 15-55; 170-215; 45-85.

(44) Page 47, lines 19-27:

The following regions in gag (SEQ ID NO:56) were determined to be antigenic by Jameson-Wolf algorithm: amino acids: 1-40; 45-60; 80-105; 130-145; 147-183; 186-220; 245-253; 255-288. Additional epitope-containing fragments include amino acids 1-8; 2-10; 1-15; 5-15; 7-15; 10-20; 12-20; 15-23; 20-28; 28-35; 30-37; 33-40; 1-20; 20-40; 1-15; 15-30; 15-40; 45-52; 50-57; 55-62; 50-60; 1-60; 80-87; 85-92; 80-90; 90-97; 95-102; 98-105; 85-100; 90-105; 80-100; 85-105; 130-137; 135-142; 140-147; 145-152; 150-157; 155-162; 160-167; 165-172; 170-177; 175-183; 180-187; 185-192; 190-197; 195-202; 200-207; 205-212; 210-217; 213-220; 185-220; 190-220; 195-220; 200-220; 205-220; 255-262; 260-267; 265-272; 270-277; 275-282; 280-288; 245-288; 250-288; 260-288; 265-288; 270-288.

(45) Page 47, line 28, to page 48, line 9:

The following regions in gag (SEQ ID NO:57) were determined to be antigenic by Jameson-Wolf algorithm: amino acids: 1-40; 80-105; 145-180; 185-225; 240-335. Additional epitope-containing fragments include amino acids 1-8; 2-10; 1-15; 5-15; 7-15; 10-20; 12-20; 15-23; 20-28; 28-35; 30-37; 33-40; 1-20; 20-40; 1-15; 15-30; 15-40; 80-87; 85-92; 80-90; 90-97;

95-102; 98-105; 85-100; 90-105; 80-100; 85-105; 145-152; 150-157; 155-162; 160-167; 165-172; 170-177; 175-182; 180-187; 185-192; 190-197; 195-202; 200-207; 205-212; 210-217; 215-212; 218-225; 145-160; 150-165; 155-170; 160-175; 170-185; 180-225; 185-225; 190-225; 195-225; 200-225; 205-225; 210-225; 215-225; 240-247; 245-252; 250-257; 255-262; 260-267; 265-272; 270-277; 275-282; 280-287; 285-292; 290-297; 295-302; 300-307; 305-312; 310-317; 315-322; 320-327; 325-332; 328-335; 245-285; 250-285; 260-285; 265-285; 270-295; 275-300; 280-305; 285-310; 295-315; 300-320; 305-325; 325-335; 245-335; 250-335; 255-335; 260-335; 270-335; 275-335; 280-335; 285-335; 290-335; 295-335; 305-335; 310-335; 315-335; 320-335.

(46) Page 48, lines 11-16:

The invention also provides an isolated polypeptide having formula 5'-A-B-C-3', wherein: A is an amino acid sequence consisting of *a* amino acids; B is an amino acid sequence consisting of a fragment of *b* amino acids from (i) the amino acid sequence encoded by a polynucleotide shown in SEQ ID NO:11, 12 or 13; (ii) any one of SEQ ~~IDs~~ ID NOS:46-49, 50-55, 56-57 or 58; C is an amino acid sequence consisting of *c* amino acids; and wherein said polypeptide is not a fragment of the amino acid sequence defined in (i) or (ii).

(47) Page 48, line 25, to page 49, line 2:

The invention provides a polypeptide having at least *s*% identity to: (a) the polypeptide sequences encoded by SEQ ~~IDs~~ ID NOS:7-45; (b) a fragment of *x* amino acids of the polypeptide sequences encoded by SEQ ~~IDs~~ ID NOS:7-45; (c) the polypeptide sequences SEQ ~~IDs~~ ID NOS:46-58; (d) a fragment of *x* amino acids of the polypeptide sequences SEQ ~~IDs~~ ID NOS:46-58. The value of *s* is at least 35 (*e.g.* at least 40, 45, 50, 55, 60, 65, 70, 75, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.5, 99.9 *etc.*). The value of *x* is at least 7 (*e.g.* 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100.

(48) Page 49, lines 13-20:

The invention provides polypeptides, such as those shown in SEQ ~~IDs~~ ID NOS:46-58, encoded by HERV-K(CH) polynucleotides that are differentially expressed in prostate cancer cells. Such polypeptides are referred to herein as "polypeptides associated with prostate cancer" or "HERV-K(CH) polypeptides". The polypeptides can be used to generate antibodies specific for a polypeptide associated with prostate cancer, which antibodies are in turn useful in

diagnostic methods, prognostic methods, therametric methods, and the like as discussed in more detail herein. Polypeptides are also useful as targets for therapeutic intervention, as discussed in more detail herein.

(49) Page 49, line 27, to page 50, line 5:

Polypeptides, such as polypeptides of the gag regions or polypeptides of the pol regions, encoded by the polynucleotides disclosed herein, such as polynucleotides having the sequences as shown in SEQ ~~ID s~~ ID NOS:7-10 and SEQ ~~ID s~~ ID NOS:14-39, and in isolates deposited with the ATCC and having ATCC accession numbers given in Table 7 and/or their corresponding full length genes, can be used to screen peptide libraries to identify binding partners, such as receptors, from among the encoded polypeptides. Peptide libraries can be synthesized according to methods known in the art (*e.g.* see refs. & ).

(50) Page 58, lines 15-24:

In one preferred embodiment of the present invention, an array comprises at least two polynucleotides, each having a sequence selected from the group consisting of SEQ ~~ID s~~ ID NOS:14-39 and polynucleotides present in isolates deposited with the ATCC and having ATCC accession numbers PTA-2561, PTA-2572, PTA-2566, PTA-2571, PTA-2562, PTA-2573, PTA-2560, PTA-2565, PTA-2568, PTA-2564, PTA-2569, PTA-2567, PTA-2559, PTA-2563, PTA-2570. In another preferred embodiment, an array comprises at least one polynucleotide having a sequence selected from the group consisting of SEQ ~~ID s~~ ID NOS:14-39 and polynucleotides present in isolates deposited with the ATCC and having ATCC accession numbers PTA-2561, PTA-2572, PTA-2566, PTA-2571, PTA-2562, PTA-2573, PTA-2560, PTA-2565, PTA-2568, PTA-2564, PTA-2569, PTA-2567, PTA-2559, PTA-2563, PTA-2570 and at least one of a polynucleotide having a sequence shown in SEQ ID NO:42 or 43.

(51) Page 64, lines 5-19:

This invention also provides methods for detecting cancer associated with elevated levels of HERV-K(CH) polynucleotides, in particular in prostate cancer, by means of (i) detecting polynucleotides having at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90% at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or at least 100% identity to the polynucleotide shown in SEQ ~~ID s~~ ID NOS:7-10 or to polynucleotides in isolates deposited with the ATCC and having

ATCC deposit accession numbers PTA-2561, PTA-2572, PTA-2566, PTA-2571, PTA-2562, PTA-2573, PTA-2560, PTA-2565, PTA-2568, PTA-2564, PTA-2569, PTA-2567, PTA-2559, PTA-2563, PTA-2570, as measured by the alignment program GCG Gap (Suite Version 10.1) using the default parameters: open gap = 3 and extend gap = 1 or polynucleotides hybridizing under high stringency conditions to the polynucleotide shown in SEQ ~~IDs~~ ID NOS:7-10; (ii) detecting polypeptides, or fragments thereof encoded by the sequences of (i); and (iii) detecting antibodies specific for one or more of the polypeptides. Furthermore, (iv) detecting particles associated with overexpression of HERV-K(CH) polynucleotides may also be used in the diagnosis of cancer, in particular, prostate cancer, and monitoring its progression.

(52) Page 65, lines 17-22:

Accordingly, the present invention provides kits for detecting prostate cancer comprising at least one of polynucleotides having the sequence as shown in SEQ ~~IDs~~ ID NOS:7-10, SEQ ~~IDs~~ ID NOS:14-39, or fragments thereof, or having the sequence found in an isolate deposited with the ATCC and having ATCC accession numbers PTA-2561, PTA-2572, PTA-2566, PTA-2571, PTA-2562, PTA-2573, PTA-2560, PTA-2565, PTA-2568, PTA-2564, PTA-2569, PTA-2567, PTA-2559, PTA-2563, PTA-2570 or fragments thereof.

(53) Page 68, lines 1-13:

Polynucleotide generally comprising at least 10 nt, at least 12nt or at least 15 contiguous nucleotides of a polynucleotide provided herein, such as, for example, those having the sequence as depicted in SEQ ~~IDs~~ ID NOS:7-10, and 3-28, are used for a variety of purposes, such as probes for detection of and/or measurement of, transcription levels of a polynucleotide that is differentially expressed in a prostate cancer cell. A probe that hybridizes specifically to a polynucleotide disclosed herein should provide a detection signal at least 5-, 10-, or 20-fold higher than the background hybridization provided with other unrelated sequences. It should be noted that "probe" as used herein is meant to refer to a polynucleotide sequence used to detect a differentially expressed gene product in a test sample. As will be readily appreciated by the ordinarily skilled artisan, the probe can be detectably labeled and contacted with, for example, an array comprising immobilized polynucleotides obtained from a test sample (*e.g.* mRNA). Alternatively, the probe can be immobilized on an array and the test sample detectably labeled.

These and other variations of the methods of the invention are well within the skill in the art and are within the scope of the invention.

(54) Page 71, lines 20-23:

Figure 6 shows an alignment of env genomic DNA sequences from 27 HERV-K viruses. A consensus sequence (SEQ ID NO:157) is shown on the bottom line.

Figures 7-9 show alignments of inferred polypeptide sequences for gag (7), pol (8) and env (9) from various HERV-K viruses, together with consensus sequences (SEQ ~~ID~~s ID NOS:158-160).

(55) Page 77, lines 4-12:

The 16 isolates were initially determined in a first pass sequencing reaction to have the sequences as shown in SEQ ~~ID~~s ID NOS:27-39, inclusive. The isolate from the normal prostate tissue was initially determined in a first pass sequencing reaction to have the sequence as shown in SEQ ID NO:41. A first pass sequencing reaction refers to a high-throughput process, where PCR reactions generate the sequencing template then sequencing is performed with one of the PCR primers, in a single direction. A search of public databases revealed that these 16 isolates have some degree of identity to regions of the human endogenous retrovirus HERV-K(II) sequence disclosed in Genbank accession number AB047240 and shown in SEQ ID NO:44, and also to HERV-K(10), but are nonetheless unique.

(56) Page 77, lines 13-19:

The isolates were subjected to a second round of nucleic acid sequencing and were found to have the sequences as shown in SEQ ~~ID~~s ID NOS:14-26, inclusive. The isolate from the normal prostate tissue was subjected to a second round of nucleic acid sequencing and found to have the sequence as shown in SEQ ID NO:40. This second round of sequencing is a customized process, where sequencing is performed on purified dsDNA template in a DNA vector. Sequencing is done from both ends of the template, forward and reverse, with primers designed from the flanking regions of the vector, and new primers are synthesized for every additional reaction needed to span the entire insert.

(57) Page 77, line 28, to page 78, line 8:

Composite HERV-K(CH) polynucleotide sequences are shown in SEQ ~~ID~~s ID NOS:7, 8, 9 and 10 and Figure 1 provides a schematic illustration of a human endogenous retrovirus and

the HERV-K(CH) species within the schematic illustration. SEQ ID NO:7 is a composite sequence of the polynucleotides SEQ ~~IDs~~ ID NOS:14-16, inclusive, and has a consensus sequence as shown in SEQ ID NO:11. This region corresponds to the gag region of a human endogenous retrovirus. SEQ ~~IDs~~ ID NOS:8 and 9 are composites sequence of the polynucleotides having a sequence as shown in SEQ ~~IDs~~ ID NOS:17-20, inclusive, and has a consensus sequence as shown in SEQ ID NO:12. This region corresponds to the 5' pol region of a human endogenous retrovirus. SEQ ID NO:10 is a composite sequence of the polynucleotides having a sequence as shown in SEQ ~~IDs~~ ID NOS:21-26, inclusive, and has a consensus sequence as shown in SEQ ID NO:13. This region corresponds to the 3' pol region of a human endogenous retrovirus

(58) Page 78, lines 18-29:

Consensus polynucleotide sequences SEQ ~~IDs~~ ID NOS:11-13 were generated with Multiple Sequence Alignment (MSA), a web implementation of the GCG Pileup and Pretty programs. The program uses a clustering algorithm similar to the Clustal program described in reference . The default values for the alignments and consensus extraction were 8 for gap open and 2 for gap extension. The poling plurality or minimum number of like sequences specified to assign a residue to the consensus sequence was 2.

(59) Page 78, lines 24-29:

The polynucleotide sequences shown in SEQ ~~IDs~~ ID NOS:14-16, inclusive, were used for the consensus polynucleotide sequence shown in SEQ ID NO:11. The polynucleotide sequences shown in SEQ ~~IDs~~ ID NOS:17-20, inclusive, were used for the consensus polynucleotide sequence shown in SEQ ID NO:12. The polynucleotide sequences shown in SEQ ~~IDs~~ ID NOS:21-26, inclusive, were used for the consensus polynucleotide shown in SEQ ID NO:13. The "N" represents where there is no qualifying minimum representative base. i.e. at least two sequences with the same base at that site.

(60) Page 79, lines 1-5:

Northern blotting of prostate cancer cell lines using nucleotides 243-end of SEQ ID NO:150 labeled as a probe indicates that they express PCAV transcripts of several sizes, corresponding to both full-length viral genomic sequences and to sub-genomic spliced transcripts

(Figure 5). Expression of such transcripts have also been observed in teratocarcinoma cell lines [15], as shown in lane 1 of figure 14.